Assignment 4

1. I implemented the algorithm 10.1 Episodic SGD to optimize the q and here below is the result for 50 episodes of Cartpole v1 problem:

A graph of a graph

Description automatically generated with medium confidence

As you can see, by using the episodic SGD, the score is climbing around 20th episode, however the stability of the model is not very stable since RL is not easy to stable with stochastic gradient descent. However the overall trend is rising up for the 50 episodic updates.

Here below is the summary of the NN structure:

Model: "sequential\_2"

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

dense\_6 (Dense) (None, 24) 120

dense\_7 (Dense) (None, 24) 600

dense\_8 (Dense) (None, 2) 50

=================================================================

Total params: 770

Trainable params: 770

Non-trainable params: 0